Developing Improved Models of Signal Transduction Pathways via Systems Biology

Juergen Hahn

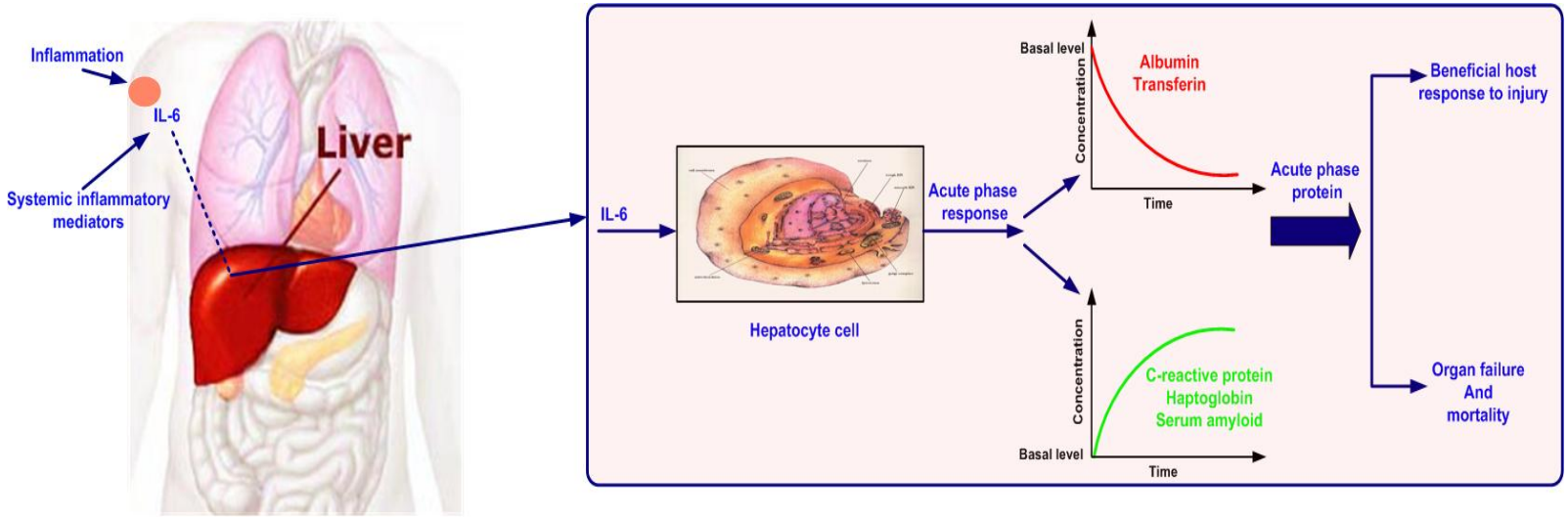
Artie McFerrin Department of Chemical Engineering
Texas A&M University
College Station, TX 77843-3122

Overview

- Introduction
- Regulatory mechanism in signaling pathways stimulated by interleukin-6
- Sensitivity analysis & model refinement
- Introduction of quantitative measurement technique for transcription factor concentrations
- Conclusions

Introduction

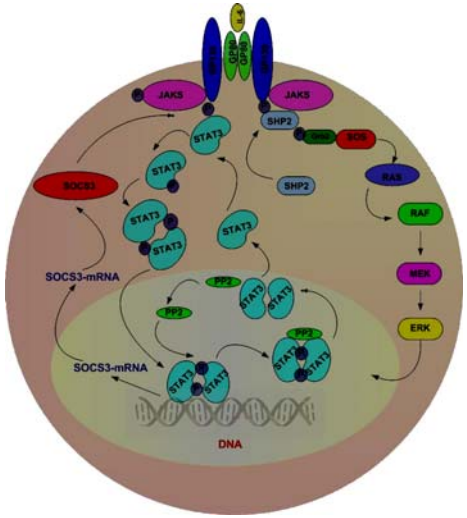
- Systemic inflammatory mediators, e.g. IL-1 and IL-6, are involved in the regulation of the hepatic acute phase response (APR)



- An improved understanding of the molecular mechanisms involved in the APR can lead to improved treatment of complications arising from inflammatory disorders

Model Description

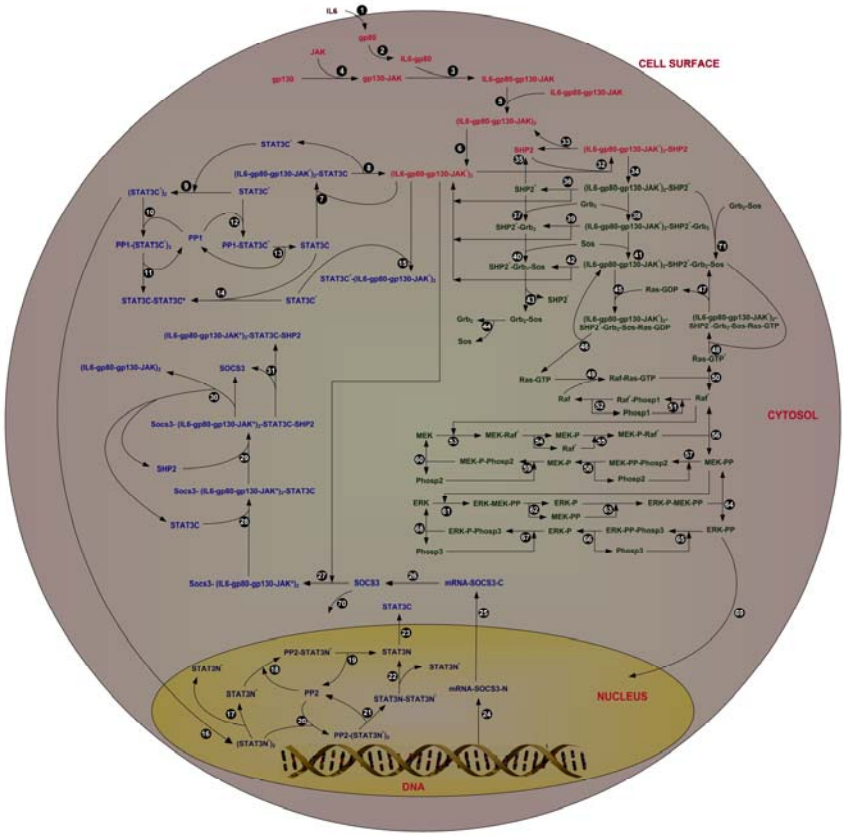
- Cell signaling induced by IL-6 in hepatocytes involves two pathways: JAK/STAT and MAPK



- The kinetic model is based on the IL-6 signaling pathway presented in Heinrich et al. (2003).
- The JAK/STAT part of model parameters have been adopted from Yamada et al. (2003) and MAPK pathway is based on the work presented by Schoeberl et al. (2002)

Yamada et al.(2003) . FEBS Letters 534: 190-196.; Schoeberl et al.(2002). Nature Biotechnology 20: 370-375.

Model Description



- Model has 68 state variables and 118 parameters
- Presented model is based on mass action/Michaelis-Menten kinetics

Singh, Jayaraman, and Hahn (2006). Biotechnology & Bioengineering 95:850-862.

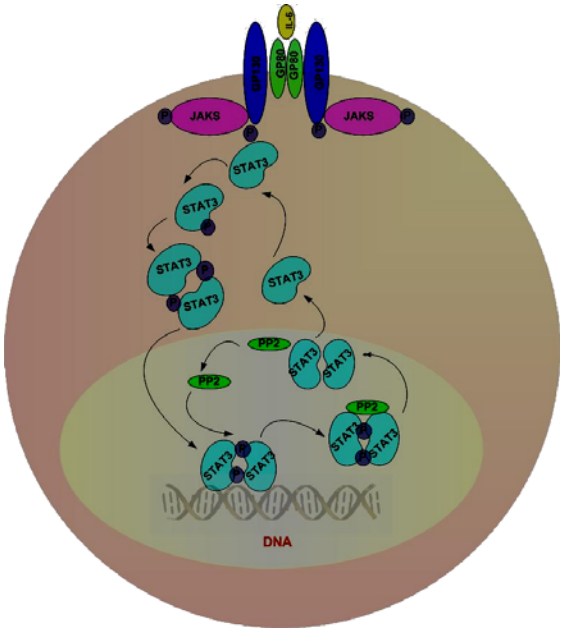
Simulation Studies

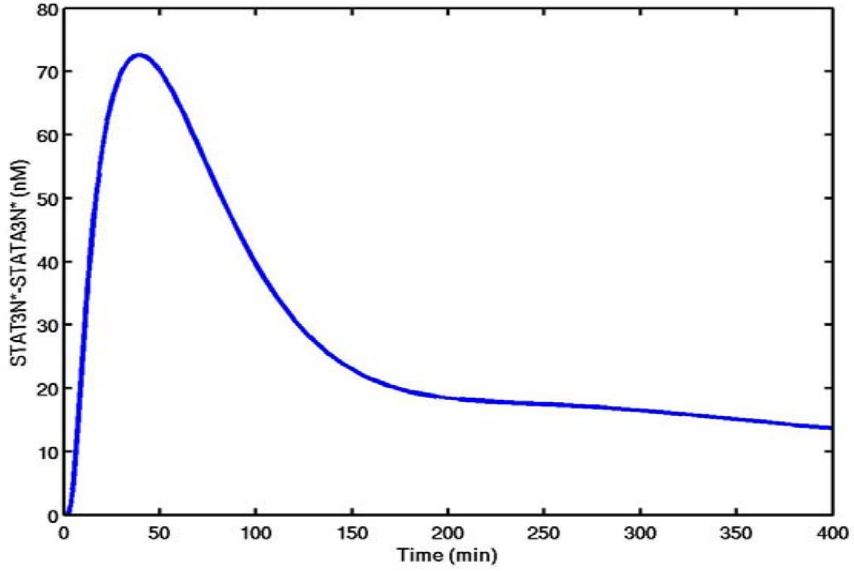
- Simulations of the developed model are compared with experimental observations in literature
- Dynamic response of signal transduction induced by IL-6
- The presented model is used to analyze the effect of:
  - SOCS3 (suppressor of cytokine signaling 3)
  - SHP-2 (domain containing tyrosine phosphatase 2)
  - STAT3 nuclear phosphatase PP2
- Further, the interactions between the two pathways has been analyzed through simulation of the developed model

Singh, Jayaraman, and Hahn (2006). Biotechnology & Bioengineering 95:850-862.

Dynamic Response

- Comparison of nuclear STAT3 dynamics with experimental results

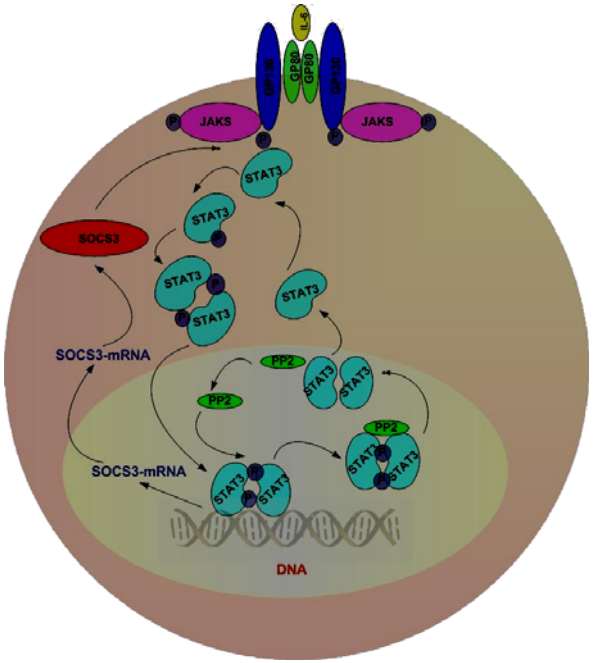


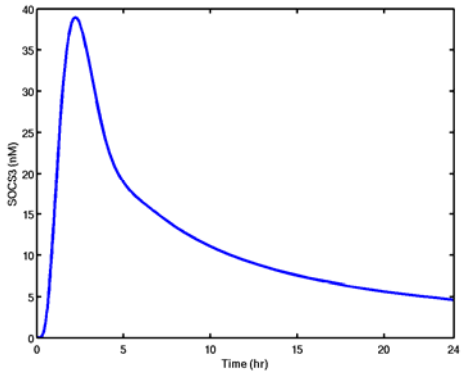


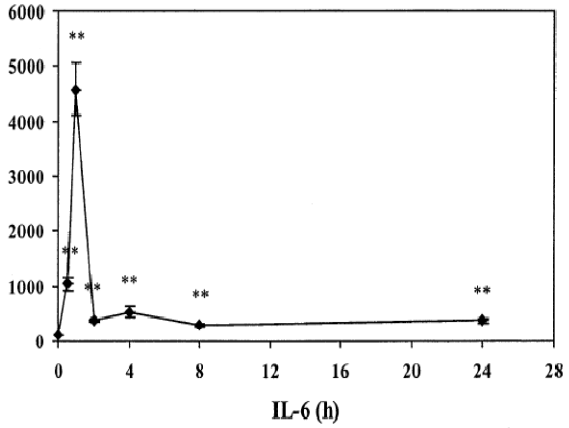
Lang et al. (2003). Nature Immunology 4: 546-550.

Dynamic Response

- SOCS3 acts as a feedback inhibitor and appears about 30 minutes after the IL-6 binding. SOCS3 reaches maximal concentration after about 2 hours



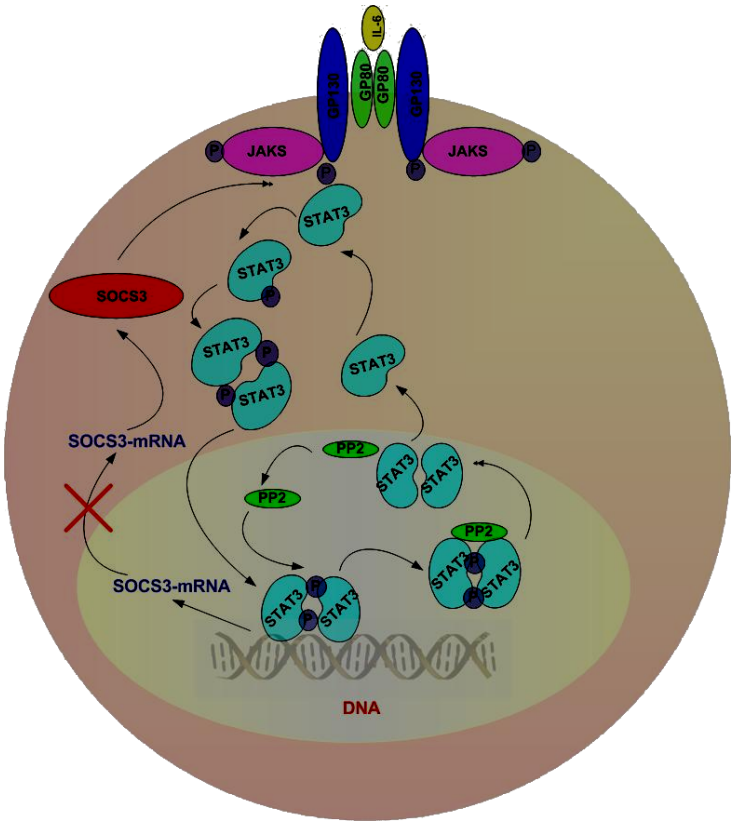


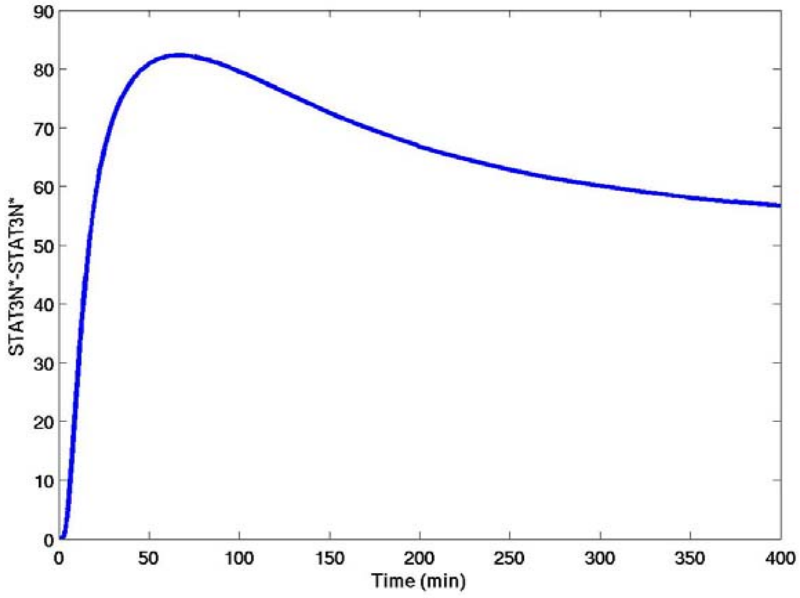


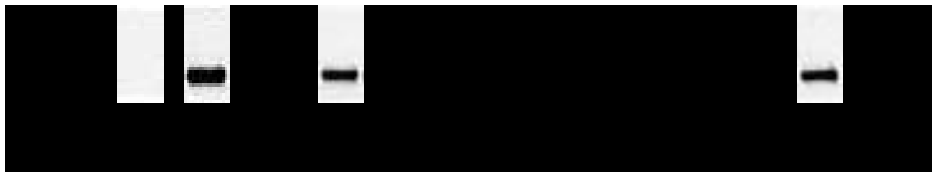
Fasshauer et al. (2004). Journal of Endocrinology 181: 129-138.

Feedback Regulation

- Influence of SOCS3 on signal transduction through the JAK/STAT pathway is investigated in a SOCS3 knockout cell



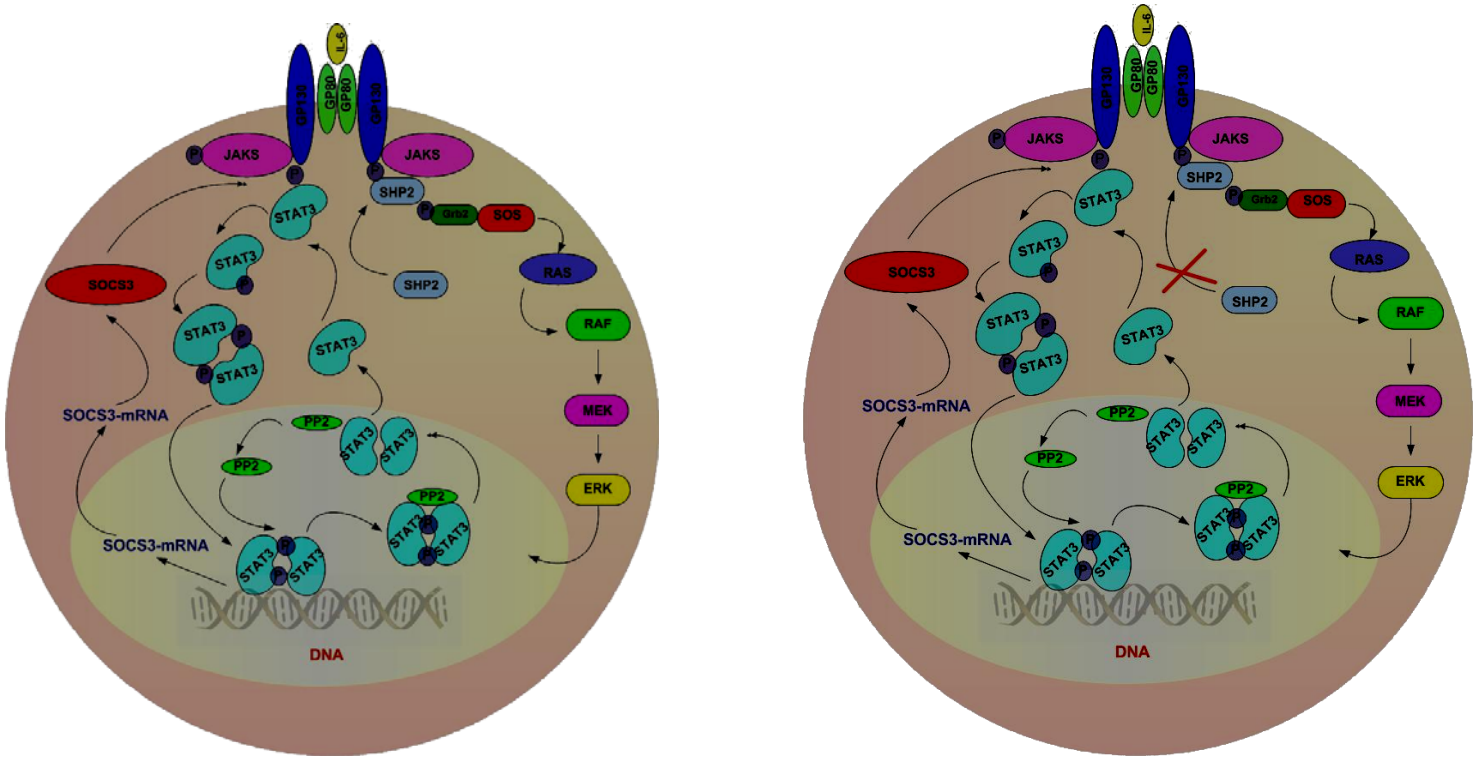




Lang et al. (2003). Nature Immunology 4: 546-550.

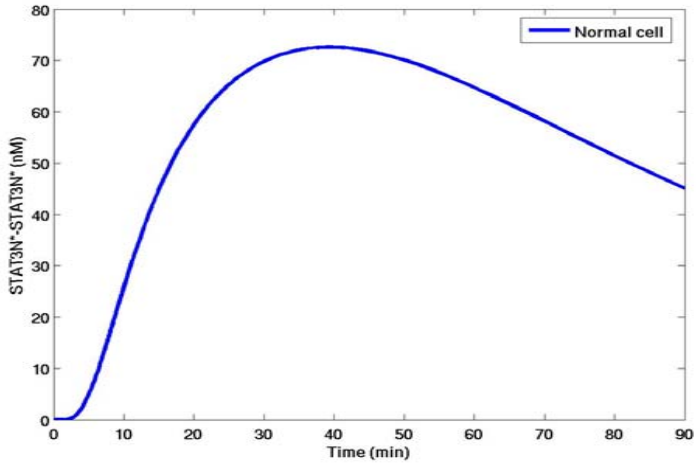
Inhibitory Effect of SHP2

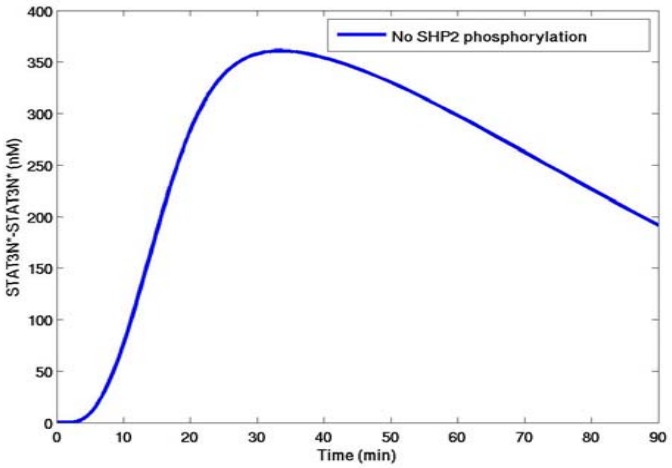
- SHP2 is not only one of the main components for initiating the MAPK pathway, it also acts as an inhibitor for the JAK/STAT pathway.



Inhibitory Effect of SHP2

- Comparison with experimental results

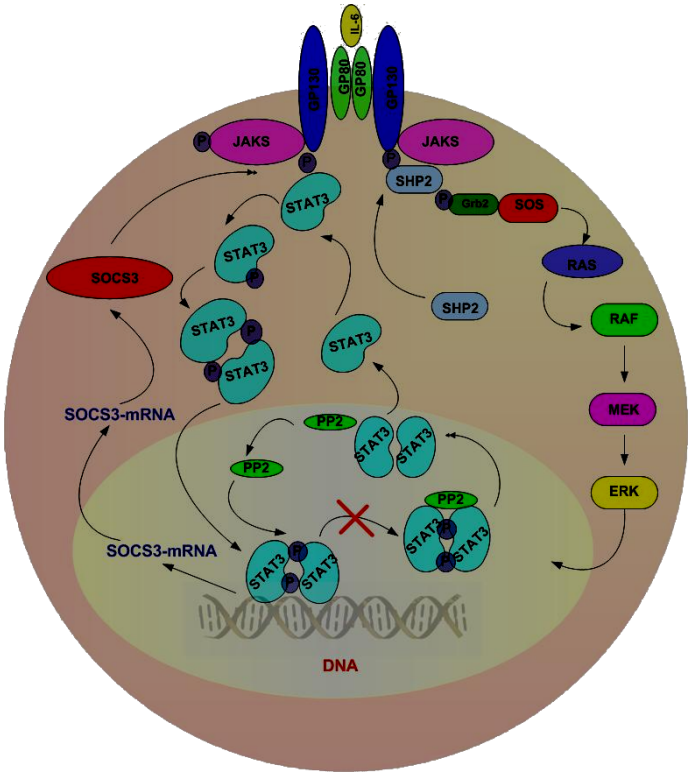


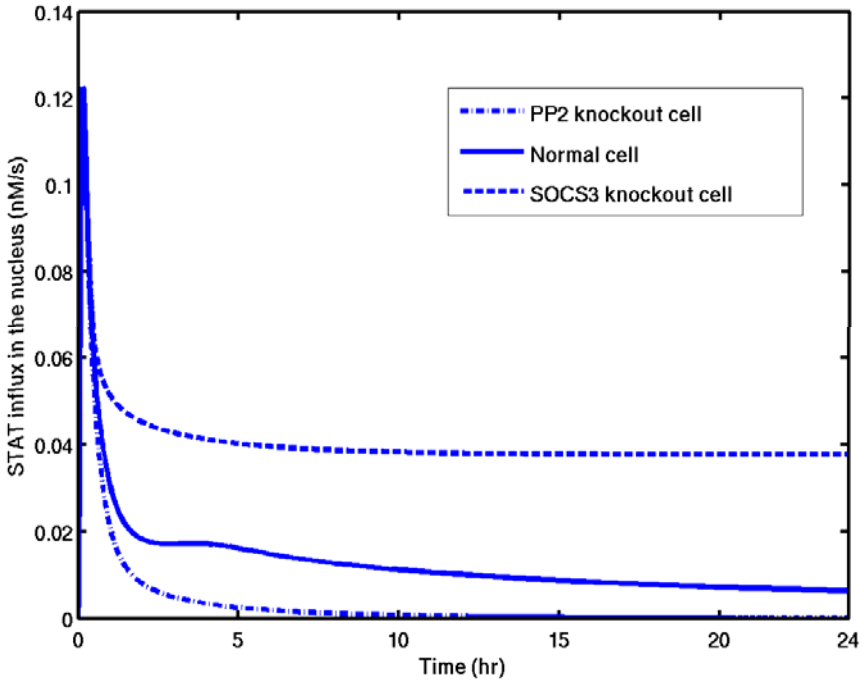


Schmitz et al. (2000). The Journal of Biological Chemistry 275: 12848-12856.

Nucleus-Cytosol Cycling of STAT3

- Nuclear export/import of STAT3 is important for signal transduction through JAK/STAT pathway
- Comparison of cell signaling in PP2 knockout, SOCS3 knockout and normal cell



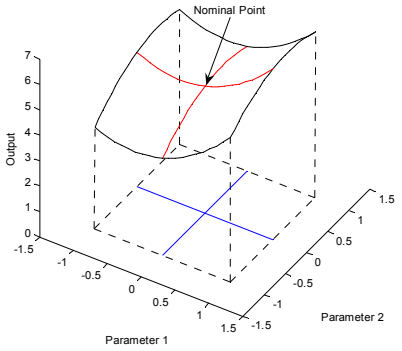


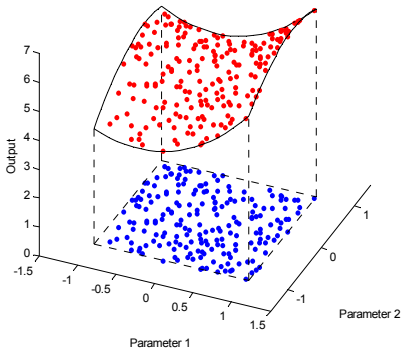
Summary of Simulation Results

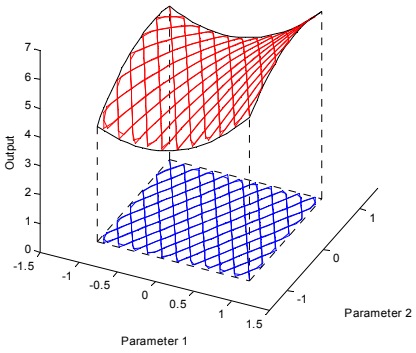
- Developed a model describing IL-6 signal transduction through the JAK/STAT and the MAPK pathways
- Model seems to describe semi-quantitative data with reasonable degree of accuracy
- SOCS3, SHP-2, and PP2 are important proteins affecting the signal transduction as shown by simulations with knockout cells
- Simulations seem to indicate significant interactions between the pathways

Parameter Sensitivity Analysis

- A more rigorous analysis than comparing simulations is required for refining the model
- Parameter sensitivity analysis is a tool which allows to quantitatively determine the affect that specific parameters have on the output
- Different approaches to parameter sensitivity analysis:
  - Local techniques (parameter output sensitivity matrix)
  - Global techniques (FAST, sampling-based approaches)







Parameter Sensitivity Analysis

- Summary of the results by the three Methods

No	Differential Analysis		Sampling-based				FAST			
			ΔP: 99-101%		ΔP: 10-1000%		ΔP: 99-101%		ΔP: 10-1000%	
	Symbol	Value	Symbol	Value	Symbol	Value	Symbol	Value	Symbol	Value
1	kf7	1	kf7	1	kf7	1	kf7	1	kf7	1
2	kf32	0.748	kf32	0.8459	kf21	0.9672	kf32	0.748	kf21	0.9782
3	kf21	0.7129	kf8	0.6992	kf8	0.7612	kf21	0.7128	Vm24	0.7639
4	kf8	0.7061	kf21	0.6519	kf26	0.7391	kf8	0.706	kf8	0.7593
5	kb7	0.6667	kb7	0.5514	Vm24	0.7114	kb7	0.6667	kf26	0.7532
6	kf20	0.5628	kf31	0.5288	kf27	0.6881	kf20	0.5627	kf27	0.7389
7	kb20	0.5492	kf48	0.5	kf20	0.6523	kb20	0.5491	kf28	0.6619
8	kf42	0.4772	kb71	0.491	kb7	0.6443	kf42	0.4773	kb7	0.6252
9	Vm24	0.4503	kb29	0.4661	kb27	0.6368	Vm24	0.4503	kf31	0.6019
10	kf26	0.4503	kf20	0.4597	kf28	0.6285	kf26	0.4503	ka26	0.5989
11	kf27	0.4472	kb28	0.4437	kf48	0.6276	kf27	0.4472	kf29	0.5936
12	kf45	0.4191	kf70	0.4109	kf70	0.5917	kf45	0.419	kf70	0.592
13	Km24	0.4131	Vm24	0.4107	kb28	0.5697	Km24	0.4132	kf48	0.5865
14	ka26	0.4077	kf42	0.3752	kb48	0.5557	ka26	0.4077	Km24	0.569
15	kf70	0.4071	kf28	0.3583	ka26	0.5324	kf70	0.4071	kf20	0.5657

Parameter Sensitivity Analysis

- Summary of the Results by the Three Methods
  - Differential sensitivity analysis returns the same results as FAST for small perturbations in the parameters
  - Set of important parameters is similar for all three methods for small perturbations
  - Set of important parameters using FAST and the sampling-based method are very similar even for large perturbations
  - Results for small and for large perturbations can be quite different, e.g., significance of kf32 is strongly influenced by the perturbation size
    - Local sensitivity analysis may not be appropriate for a system with large uncertainties
  - The most important parameters can be directly linked to terms involved in the formation and disassociation of certain proteins
    - SOCS3 (Vm24, kf26, kf27),
    - PP2 (kf20, kf21),
    - SHP-2(kf48,kf32),

    as well as the initiation of signaling through the JAK/STAT pathway (kf7)
  - kf7 seems to play a major role in activity of the JAK/STAT vs. activity of the MAPK pathway

Chu, Jayaraman, and Hahn (2007). IET Trans Systems Biology 1:342-352

Parameter Sensitivity Analysis

No.	10% kf7		25% kf7		100% kf7		400% kf7		1000% kf7	
2	kf48	22.88	kf48	45.277	kf21	127.35	kf21	323.21	kf21	511.08
3	kf32	18.081	kf21	42.556	kf32	106.94	kf8	243.69	kf8	390.52
4	kf71	16.949	kf32	39.696	kf8	99.754	kf27	232.26	kf27	384.29
5	kf21	16.548	kf71	37.665	kf48	99.245	kf26	228.37	kf26	378.72
6	kb7	15.759	kb7	35.332	kb7	95.061	Vm24	225.41	Vm24	378.13
7	kf8	14.831	kf8	34.198	kf26	89.239	kb7	218.43	kb7	328.7
8	kf20	14.648	kf20	32.326	kf20	87.890	kf32	198.67	kf31	324.98
9	kb48	14.523	Vm24	31.753	kf27	87.203	kf31	194.15	kf28	323.04
10	kf26	12.357	kf26	31.492	Vm24	85.295	kf20	191.61	kf70	307.79
11	Vm24	12.307	kb48	30.322	kf71	83.552	kf28	189.73	kf29	307.24
12	kf42	11.993	kf42	28.416	kb48	77.612	kf29	183.4	ka26	306.14
13	kf19	11.922	kf27	27.999	kf31	74.358	kf48	183.24	kf20	280.47
14	kb18	11.738	kf19	27.856	Km24	69.805	kf70	180.86	Km24	276.32
15	kf18	11.498	kb20	27.677	kf29	69.504	ka26	180.68	kb27	276.1
16	kb20	10.831	kf28	27.443	ka26	69.127	Km24	171.16	kb28	274.2
17	Kf45	9.9303	kf31	24.98	kf28	69.028	kb48	165.23	kb29	257.07
18	kf27	9.9222	kf18	24.165	kf70	67.894	kb28	162.67	kf32	242.28
19	ka26	9.8	kf29	23.371	kf42	67.321	kb27	161.49	kf48	239.7
20	kf31	9.7821	kf70	23.362	kf19	63.411	kb29	150.92	kb48	235.82
21	Km24	9.7425	Km24	22.964	kb20	62.759	kf71	149.57	kb20	213.38

Parameter Sensitivity Analysis

- The value of kf7 does have a profound impact not only on activity in the individual pathways but also on the rank of other parameters computed from sensitivity analysis
- Many parameters associated with the MAPK pathway increase in importance for small values of kf7
- Note that the output for all these investigations is STAT3 which is a “product” of the JAK/STAT pathway

  → Indicator of strong interactions among the pathways

  → Small changes in one part of the pathway can have a profound impact on the entire signaling activity
- SHP-2, Grb2, and Sos play a major role for crosstalk between pathways as output is influenced by kf48, kf71, kb48, kf42

Chu, Jayaraman, and Hahn (2007). IET Trans Systems Biology 1:342-352

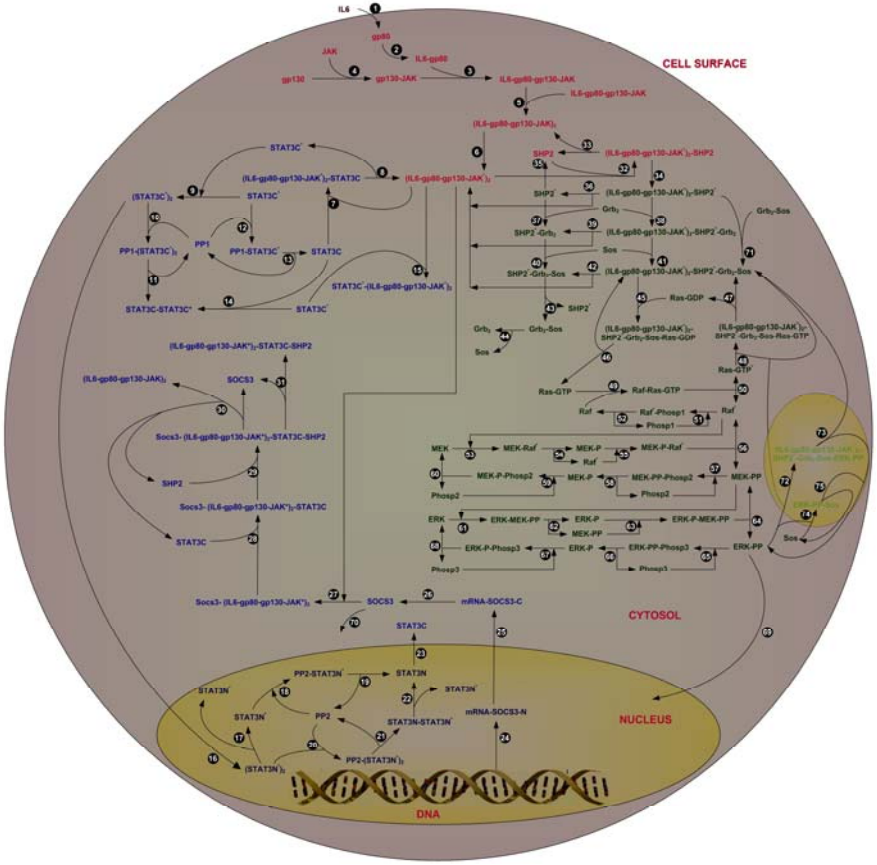
Parameter Sensitivity Analysis

- It had been suggested (Orton et al., 2005) that there is a hidden feedback loop between ERKPP and Sos which was not modeled in the model proposed by Schoeberl et al. (2002)
- Since sensitivity analysis indicated that there is significant crosstalk between pathways and that (IL6-gp80-gp130-JAK*)$_2$-SHP2*-Grb2-Sos is a key protein complex, modeling of this feedback loop can have influence of STAT3 concentration profile
- Implement loop where ERKPP phosphorylates Sos causing it to disassociate from the receptor complex

Schoeberl et al. (2002). Nature Biotechnology 20: 370-375; Orton et al. (2005). Biochem. J. 392 : 249–261.

Modeling ERKPP-Sos Feedback Loop

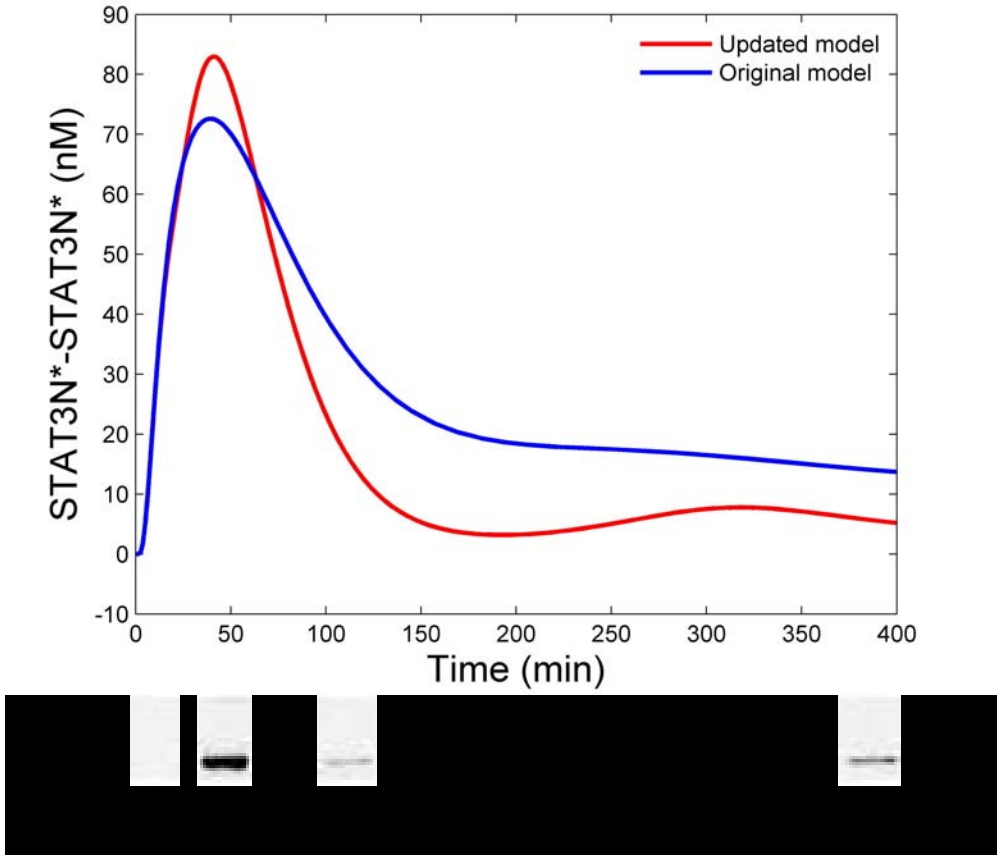
- Updated model



- New model has 72 states and 124 parameters

Modeling ERKPP-Sos Feedback Loop

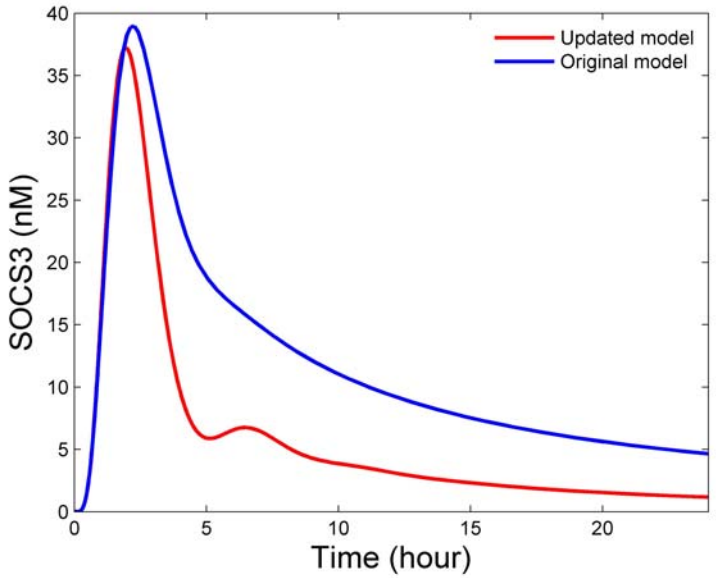
- Comparison of nuclear STAT3 for old and new model with results from the literature

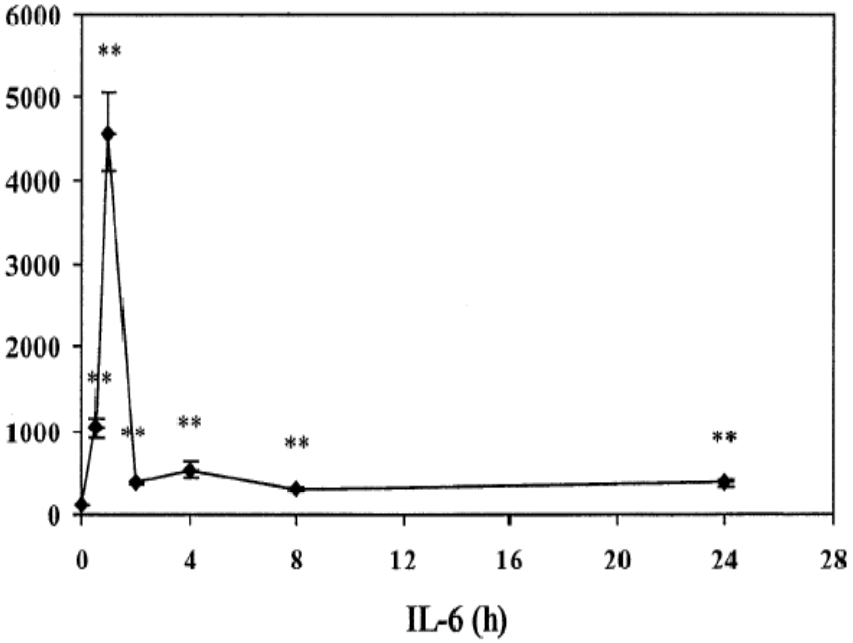


Lang et al. (2003). Nature Immunology 4: 546-550.

Modeling ERKPP-Sos Feedback Loop

- Comparison of SOCS3 dynamics for old and new model with results from the literature





Fasshauer et al. (2004). Journal of Endocrinology 181: 129-138.

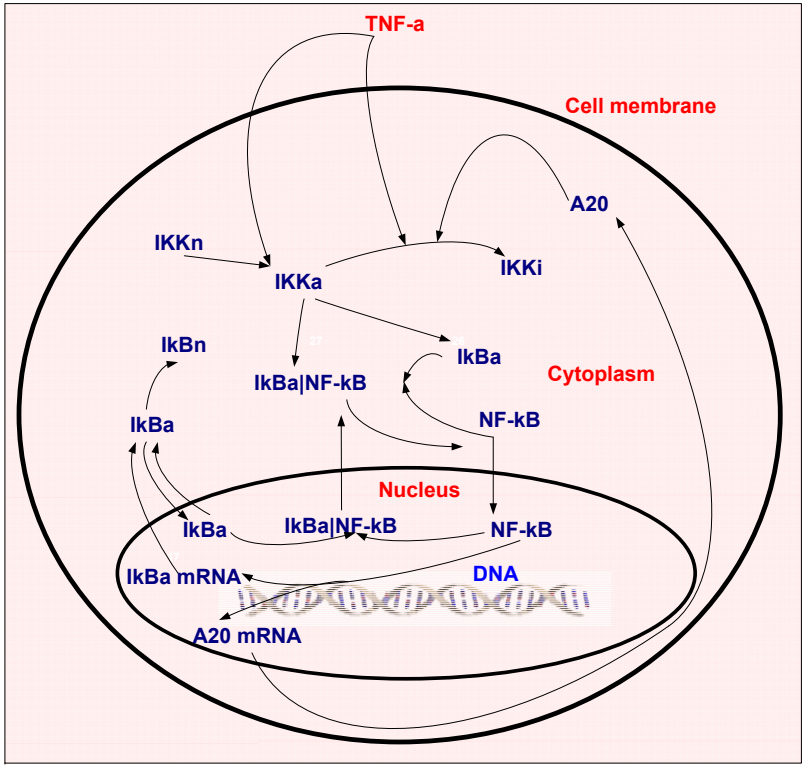
Analysis of Experimental Data and Comparison

- All experimental data shown so far came from the literature
- Potential problem:
  - Western blots provide semi-quantitative data
  - Western blots require that each sample is destroyed for taking one measurement

  → Require a new measurement technique that returns quantitative information
- Develop new measurement technique on TNF-α signaling pathway and then apply it to IL-6 signaling pathway

TNF-α Signaling Pathway

- Cell stimulated with TNF-α show increased levels of the transcription factor NF-κB associated with this pathway
- Qualitative data exists in the form of Western blots for proteins involved in this pathway[1]
- Several qualitative mathematical models have been presented[1,2,3]
- Each of the existing models only describes part of the signal transduction pathway



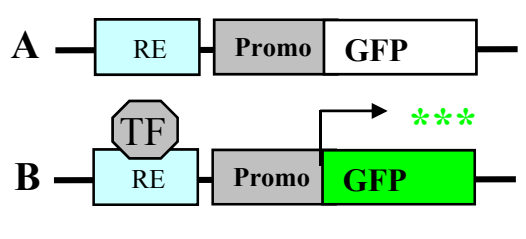
[1]A. Hoffmann, A. Levchenko, M.L. Scott, and D. Baltimore (2002) Science

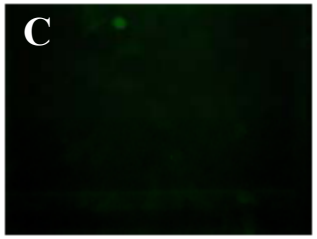
[2]P.Rangamani and L. Sirovich (2007) Biotech. Bioeng.

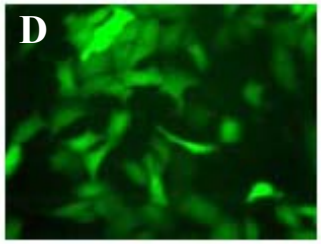
[3]T. Lipniacki, P. Paszek, A.R. Brasier, B.A. Luxon, and M. Kimmel (2004) J. Theor. Biol.

Fluorescence-based Reporter Systems

- Fluorescence-based reporter systems can be created where GFP is produced in addition to the proteins that are transcribed/translated by a particular transcription factor
- GFP expression and fluorescence is observed only when a transcription factor binds to DNA
- Compared to Western blots, GFP reporter systems allow continuous and non-invasive monitoring of transcription factors
- Challenge: to obtain quantitative concentration profiles from experiments

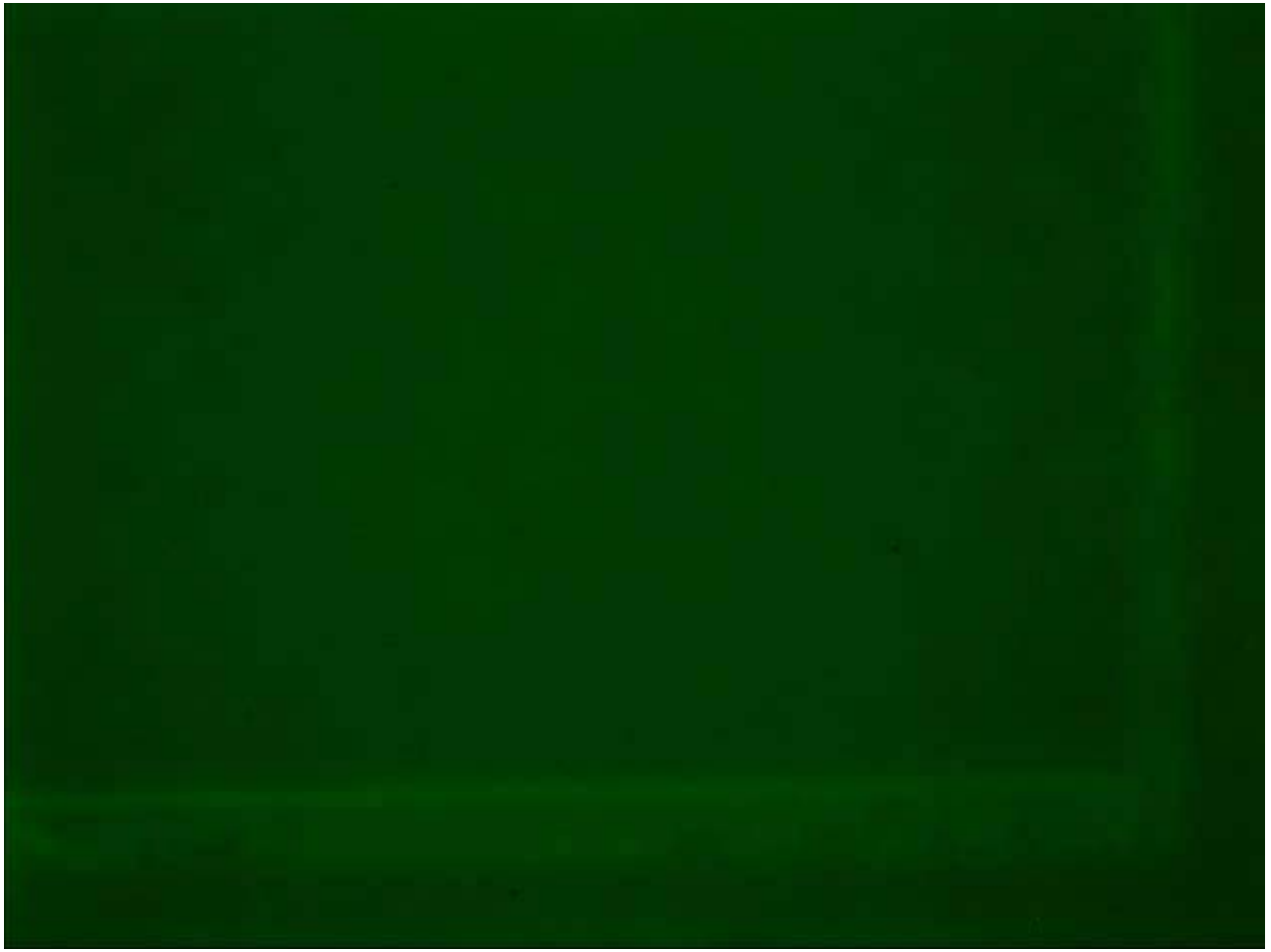






GFP-based TF reporter. (A) No fluorescence in the absence of TF binding, (B) Binding of TF leads to promoter activation and GFP fluorescence. Fluorescence images (C) without and (D) with NF-κB activation by TNF-α are shown.

Fluorescence-based Reporter Systems

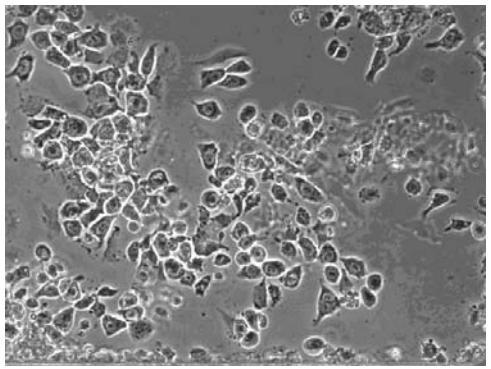


Quantitative Measurement Technique for Transcription Factor Profiles

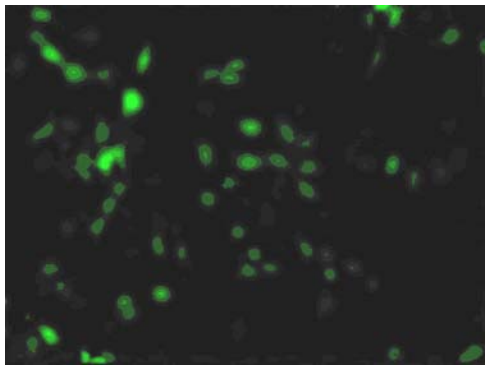
- Determining transcription factor profiles requires three tasks:
  - Generation of GFP reporter system and creation of experimental data
  - Determining fluorescence intensity profiles from fluorescence microscopy images taken at different points in time
  - Solution of an inverse problem to determine transcription factor concentration from the fluorescence intensity profile
- Focus here is on the 2nd and 3rd task

Image Analysis

- Image analysis extracts information from a time-series of images
- For each image, the following steps need to be performed:
  - Determine the area in the image representing cells where fluorescence can be seen
  - Compute the average fluorescence intensity over fluorescent area



Light microscopy image



Fluorescence microscopy image

Fluorescent cell region

Image Analysis based on K-means Clustering and PCA

- Principal component analysis can be used to distinguish regions of the image with similar features, e.g., brightness
- Form cluster for pixels with the same brightness
- Brightest clusters represent fluorescent cells
- Procedure: Compute

$$T = XP$$

  - Distance from PC1 is an indicator of variations in the cluster
  - T(:,1) represents the projection of points in the cluster onto PC1

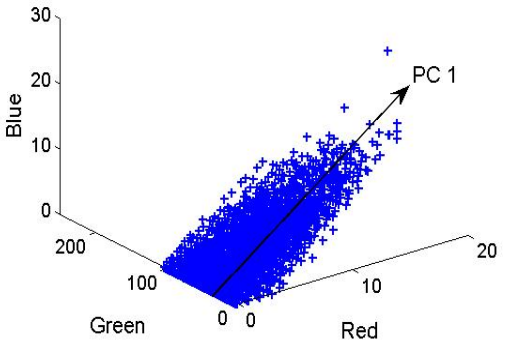


Image Analysis based on K-means Clustering and PCA

- Perform principal component analysis on images to determine centroids of clusters
- On the basis of the initial centroids of all clusters, perform K-means clustering to divide cluster
- Cluster with largest percentage of variance of PCA 1 will be selected as cluster to be divided in the next step
- Clusters with highest fluorescence intensity are considered to represent fluorescent cells, while other clusters are determined to be background

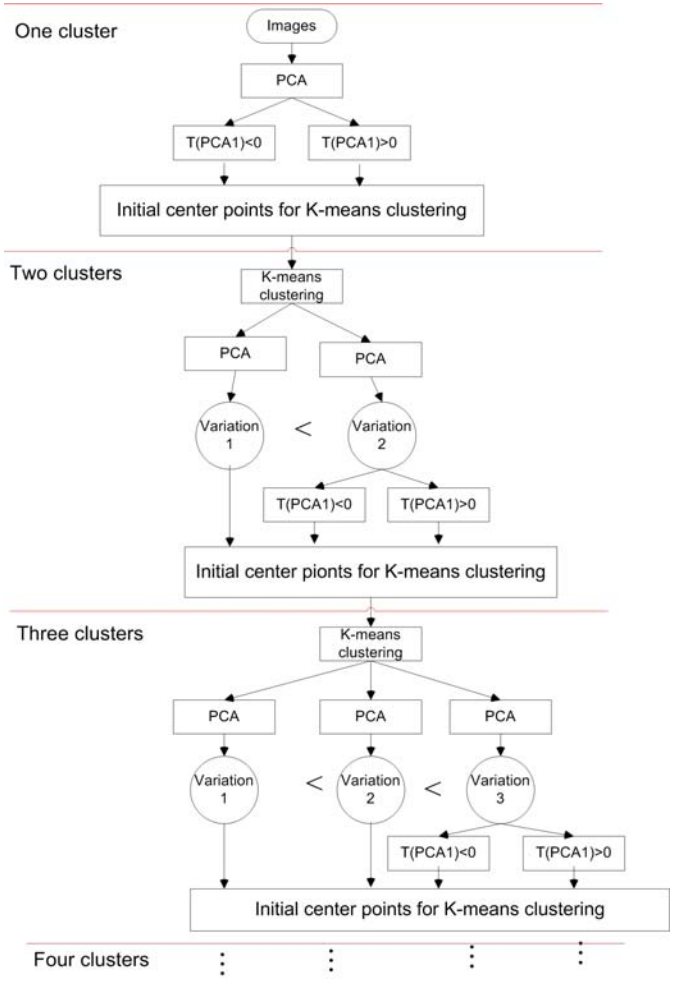
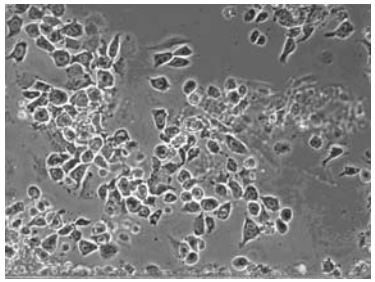
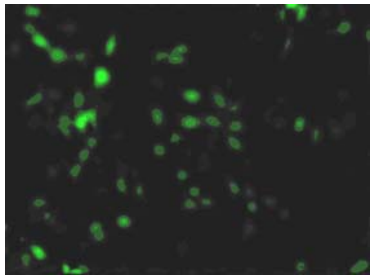


Image Analysis based on K-means Clustering and PCA

- Procedure can provide information about different intensity levels (white region indicates points belonging to a particular cluster)

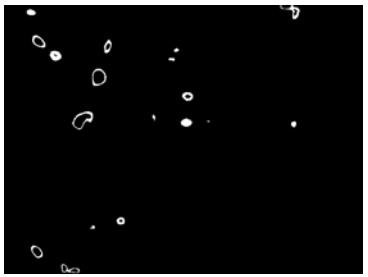


Light microscopy image

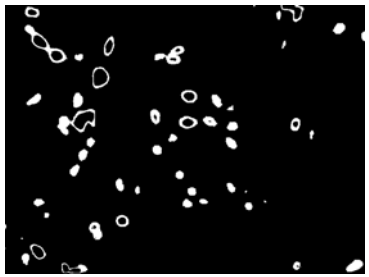


Fluorescence microscopy image

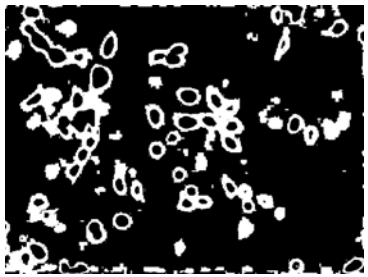
Cluster 1



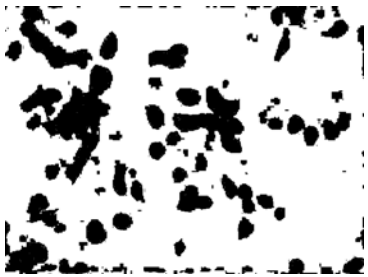
Cluster 2



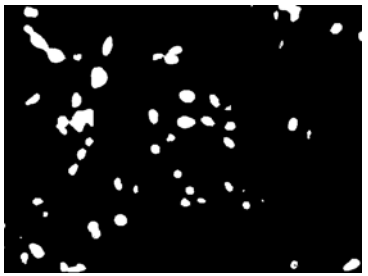
Cluster 3



Cluster 4



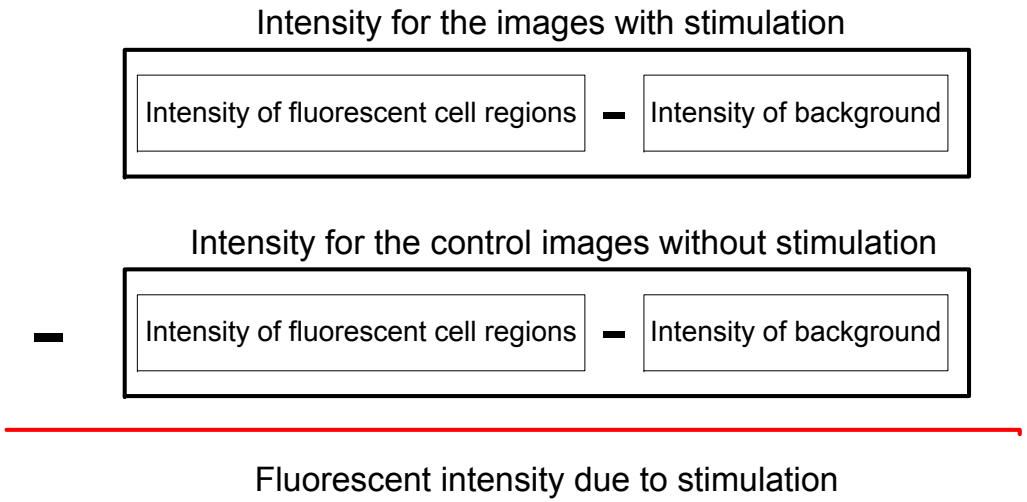
Cluster 5



Cluster 1+ Cluster 2+ Cluster 3

Computation of Fluorescence Intensity Profile

- Calculation of fluorescence intensity
  - Subtract background intensity to reduce measurement noise due to brightness variation
  - Subtract the intensity of (negative) control experiments to reduce other effects that can cause fluorescence



Computation of Fluorescence Intensity Profile

- Example of calculation of fluorescence intensity profile (H35 cells with NF-κB-GFP)

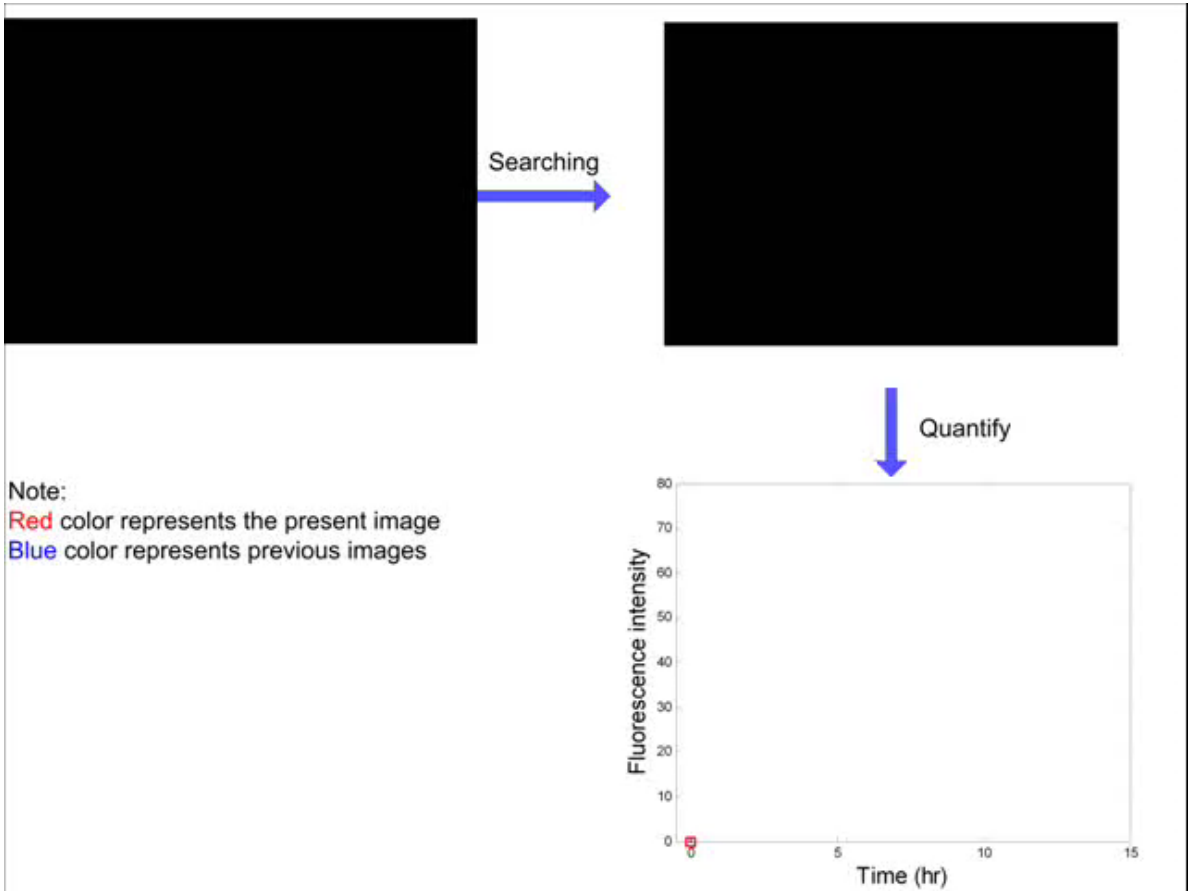
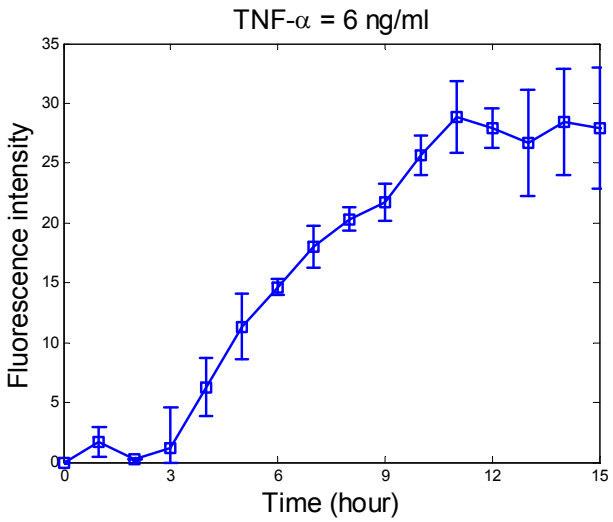
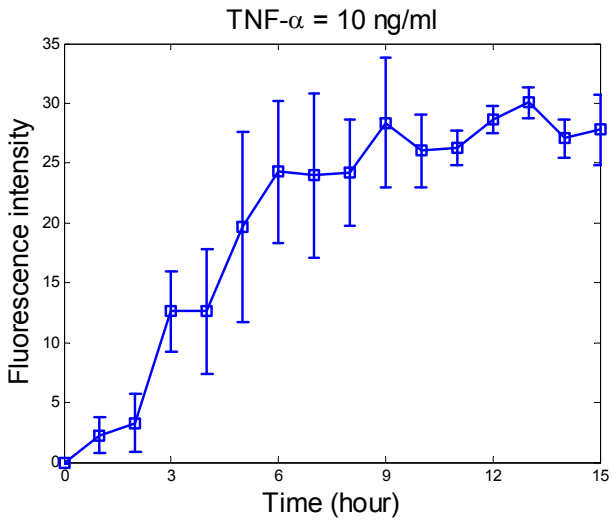
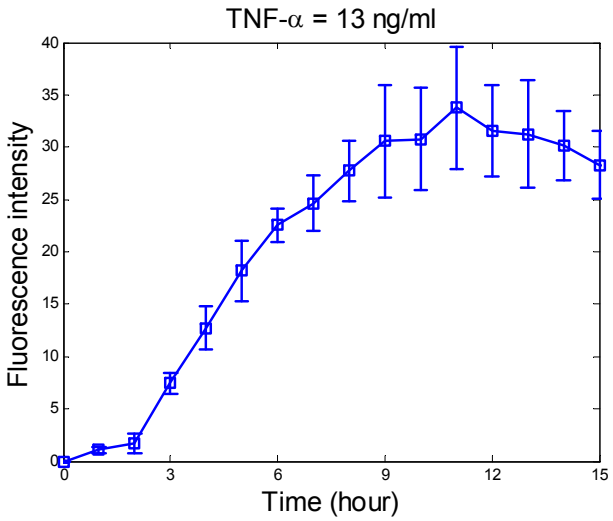
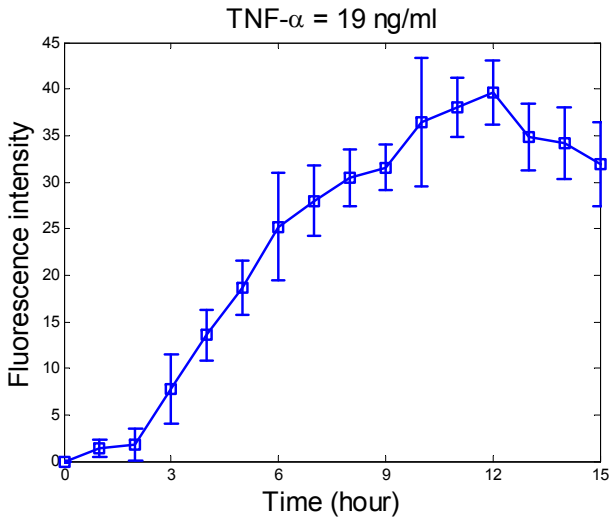


Image Analysis Applied to TNF-α Signaling Pathway





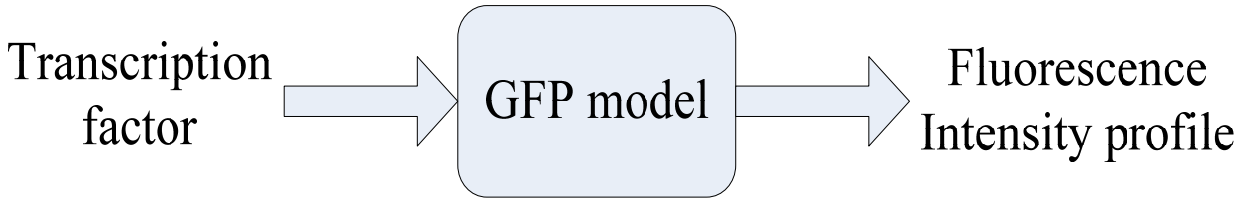




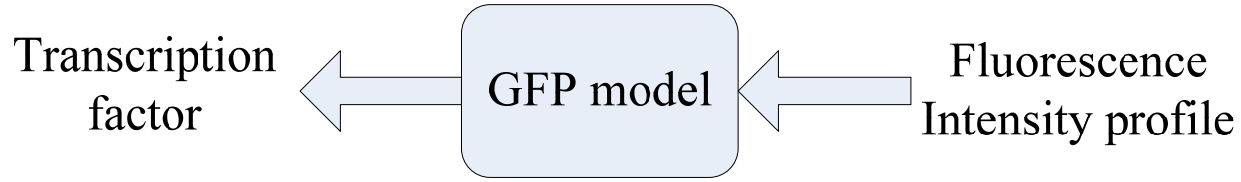
Deriving Transcription Factor Concentrations from Images

- Derive a model that correlated transcription factor concentration with fluorescence intensity
- Model needs to include transcription, translation, and activation of GFP
- Implement a procedure of system inversion to compute transcription factor profile by using this model and the fluorescence data derived from image analysis

GFP model:



System inversion:



Model for GFP Dynamics

- Modify a model for GFP dynamics taken from the literature[1] and adapt it to our experiments:
  - Amount of DNA remains constant
  - Effect of transcription factor concentrations on transcription rate is included
  - Estimate parameters C and Δ

$$dp/dt = 0$$

$$dm/dt = S_m \frac{C_{\mathrm{NF-\kappa B}}}{C + C_{\mathrm{NF-\kappa B}}} p - D_m m$$

$$dn/dt = S_n m - D_n n - S_f n$$

$$df/dt = s_f n - D_n f$$

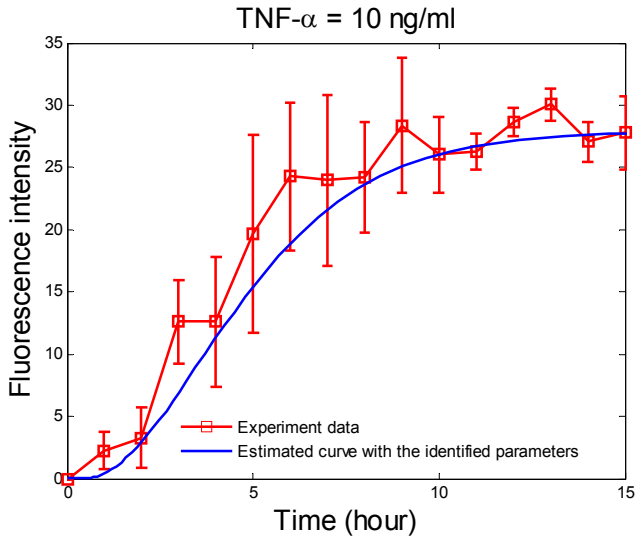
$$I = f/\Delta$$

Parameters	Variables
S_m (373 1/hr)	***p***: DNA concentration
D_m (0.45 1/hr)	***m***: GFP m-RNA , μM
S_n (780 1/hr)	***n***: non-fluorescent protein,
D_n (0.5 1/hr)	***f***: fluorescent protein
S_f (0.347 1/hr)	***I***: fluorescent intensity
Δ (estimated)	2.556×10^4
C (estimated)	108 nM

[1]Subramanian and Srienc (1996) J. Biotechnology

Model for GFP Dynamics

- Validate model by using transcription factor concentration profile taken from the literature[1] with the new model describing GFP dynamics
- Fluorescence intensity profile matches experimental observation

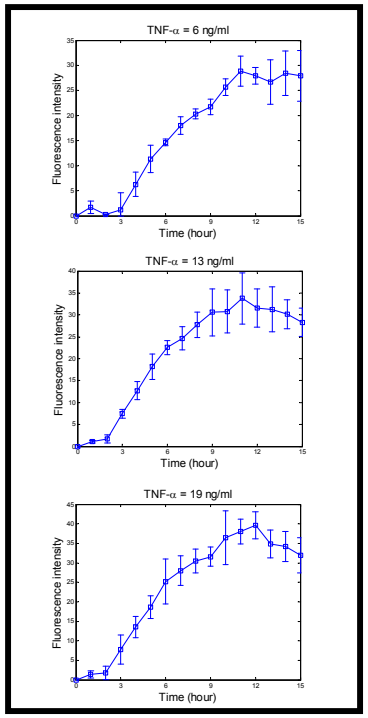


[1]A. Hoffmann, A. Levchenko, M.L. Scott, and D. Baltimore (2002) Science

Solving Inverse Problem to Obtain NF-κB Concentration

- Infer dynamics of NF-κB from fluorescence intensity data *I*

Input: $\boldsymbol{I}$



GFP model

$$dm/dt = S_m \frac{C_{\text{NF}-\kappa\text{B}}}{C + C_{\text{NF}-\kappa\text{B}}} p - D_m m$$

$$dn/dt = S_n m - D_n n - S_f n$$

$$df/dt = s_f n - D_n f$$

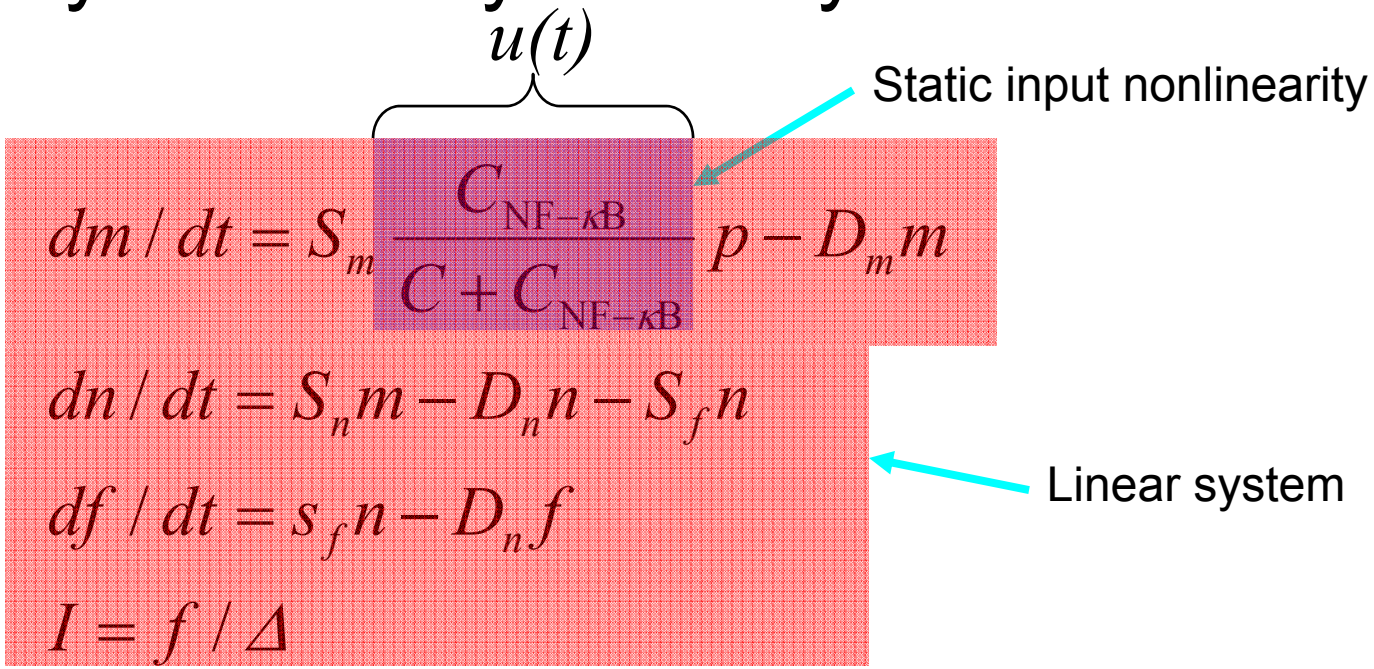
$$I = f/\Delta$$

Output: $C_{\text{NF-kB}}$

NF-kB time-series profile

Solving Inverse Problem to Obtain NF-κB Concentration

- Define $u(t) = C_{NF\text{-}\kappa B}/(C + C_{NF\text{-}\kappa B})$ such that the relationship between $u(t)$ and $I(t)$ is given by a linear dynamic system



Solving Inverse Problem to Obtain NF-κB Concentration

- $u(t)$ can be represented by a third order system to obtain the Laplace transform I(s) of I(t)

$$U(s) = \frac{\omega_n^2}{s^2 + 2\varepsilon\omega_n s + \omega_n^2} \cdot \frac{T_\alpha}{s}$$

$$\mathrm{I}(s) = \frac{1}{\Delta} \cdot \frac{S_f}{s + D_n} \cdot \frac{S_n}{s + D_n + S_f} \cdot \frac{S_m p}{s + D_m} \cdot \frac{\omega_n^2}{s^2 + 2\varepsilon\omega_n s + \omega_n^2} \cdot \frac{T_\alpha}{s}$$

- Estimate parameters ε, ω_n, T_α by fitting $I(t)$ to the fluorescence intensity data via nonlinear least squares optimization

$$\mathrm{I}(t) = A_1 + A_2 e^{-D_n t} + A_3 e^{-(D_n + S_f)t} + A_4 e^{-D_m t} + A_7 e^{-\varepsilon\omega_n t} \sin(\omega_n \sqrt{1 - \varepsilon^2}\, t + \varphi)$$

Solving Inverse Problem to Obtain NF-κB Concentration

- $C_{\text{NF-kB}}$ is then given by

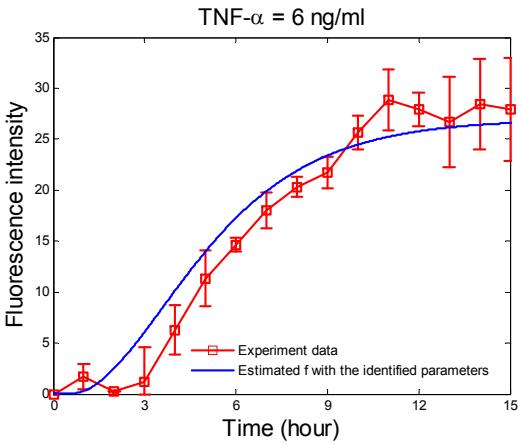
$$C_{\text{NF}-\kappa\text{B}}(t) = \frac{CT_\alpha\sqrt{1-\varepsilon^2} - CT_\alpha e^{-\varepsilon\omega_n t}\sin(\omega_n\sqrt{1-\varepsilon^2}\,t+\phi)}{(1\text{-}T_\alpha)\sqrt{1-\varepsilon^2} + T_\alpha e^{-\varepsilon\omega_n t}\sin(\omega_n\sqrt{1-\varepsilon^2}\,t+\phi)}$$

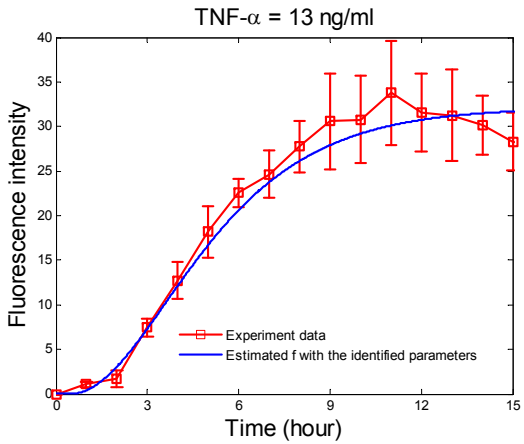
where the parameters ε, ω_n, T_α have the values that were estimated from fitting *I(t)* to the fluorescence intensity data

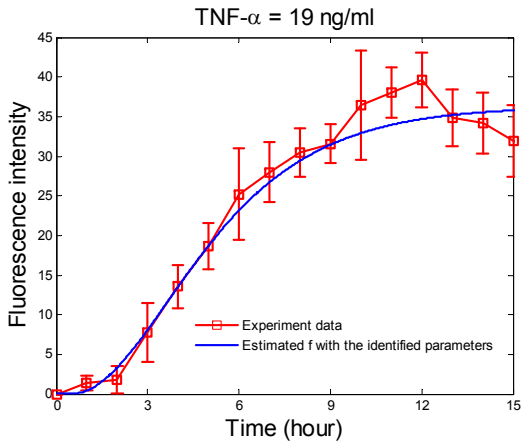
Solving Inverse Problem to Obtain NF-κB Concentration

- Estimated values for parameters $\varepsilon, \omega_n, T_\alpha$

TNF-α concentration	ε	ω_n	T_a
6 ng/ml	0.20	4.52	0.26
13 ng/ml	0.20	4.52	0.31
19 ng/ml	0.28	4.61	0.35

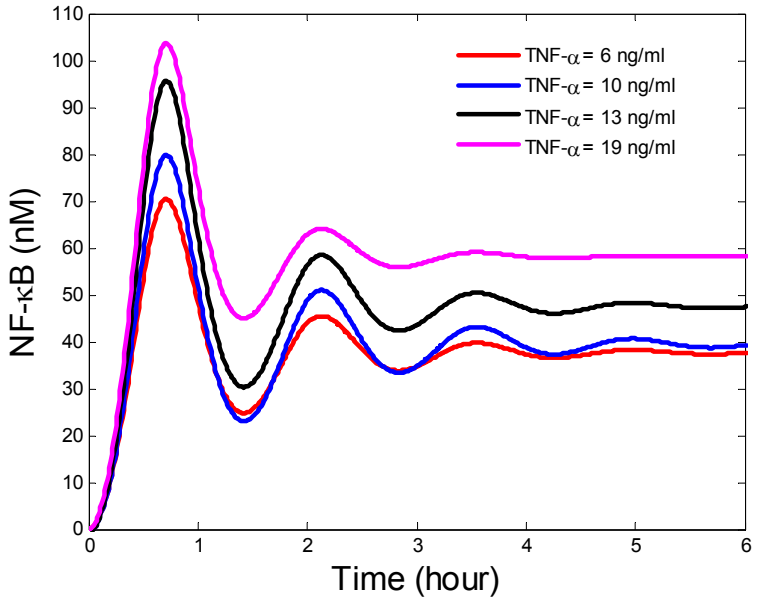






Solving Inverse Problem to Obtain NF-κB Concentration

- NF-κB profiles obtained from system inversion



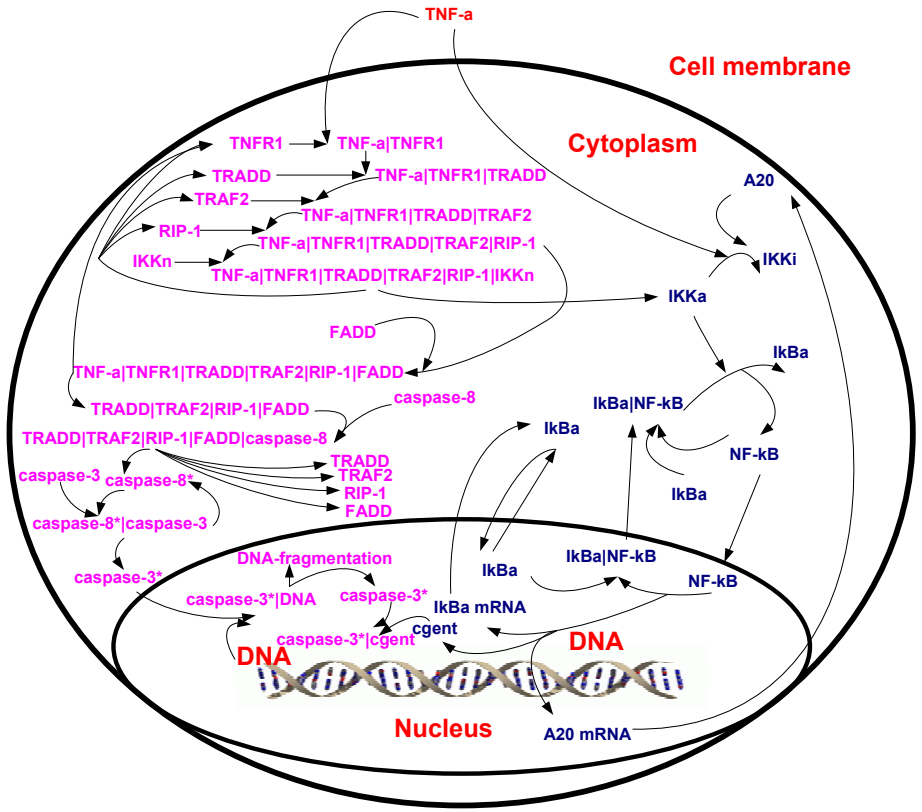
← Technique developed in this work

← Western blot data

(TNF-α = 10ng/ml)

Modeling of TNF-α Signal Transduction Pathway

- Model structure
  - Pathway from TNF-α to IKKn is taken from one source[1] and pathway from IKKn to NF-κB is taken from a different paper[2] to capture as much detail as possible



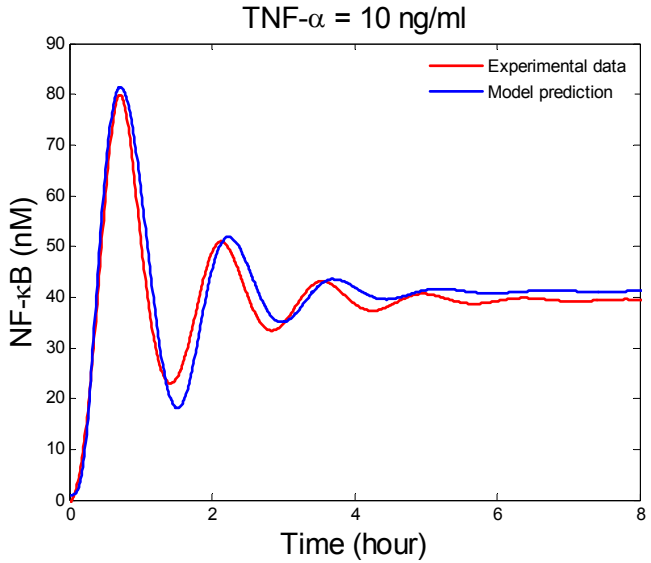
37 differential equations
57 parameters

[1]P.Rangamani and L. Sirovich (2007) Biotech. Bioeng.

[2]T. Lipniacki, P. Paszek, A.R. Brasier, B.A. Luxon, and M. Kimmel (2004) J. Theor. Biol.

Parameter Estimation for TNF-α Signaling Pathway

- Sensitivity analysis is used to select parameter C_3, k_{1p} and k_r for estimation
- Data set for 6 ng/ml, 13 ng/ml and 19 ng/ml of TNF-α are used for parameter estimation
- C_3, k_{1p} and k_r are estimated to be 0.0132, 0.0666 and 2.40, respectively
- Data set for 10 ng/ml of TNF-α is used as testing set

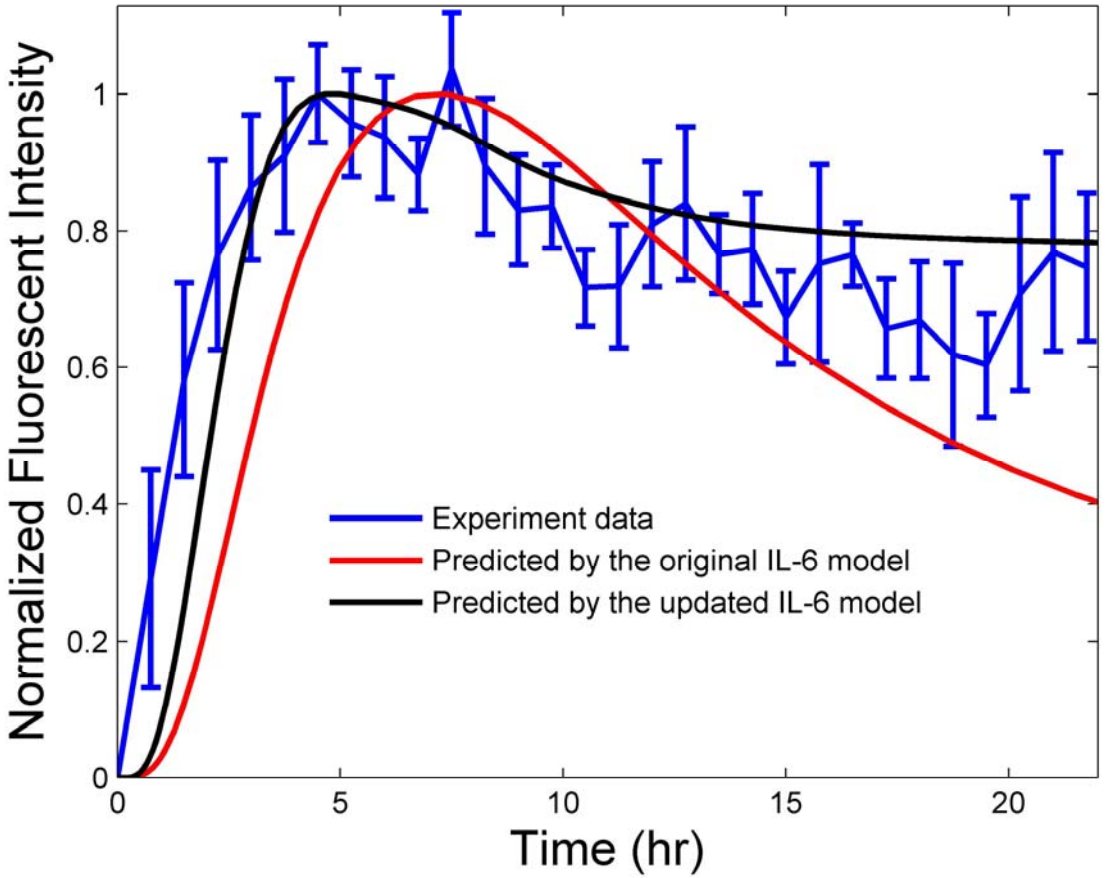


Analysis of Experimental Data and Comparison

- Experimental results are in excellent agreement with results from the literature and model predictions
- Apply measurement technique to hepatocytes stimulated by IL-6 and with STAT3 GFP reporter system
- Hepatocytes stimulated over a 8 hours time interval with IL-6
- Record images of fluorescence every 30 minutes and quantify fluorescence

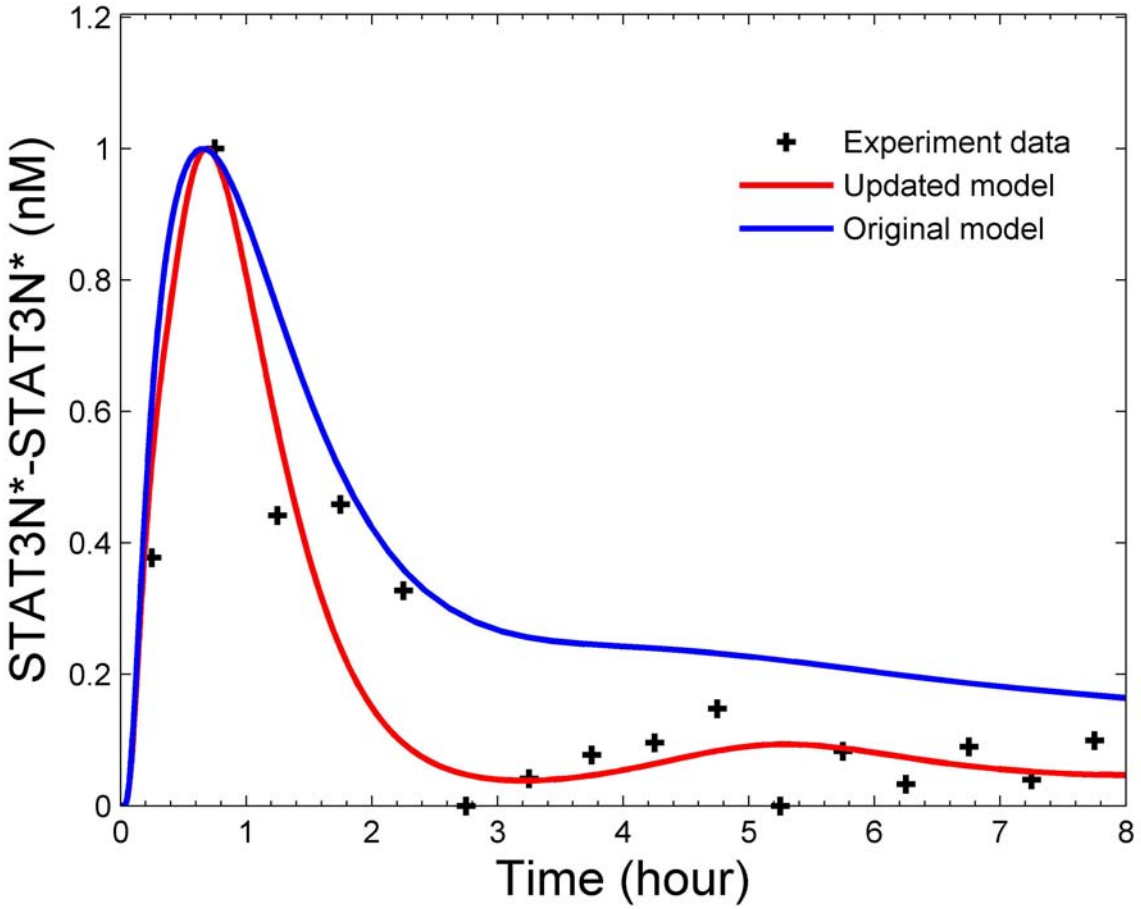
Analysis of Experimental Data and Comparison

- Results



Analysis of Experimental Data and Comparison

- Results



Conclusions

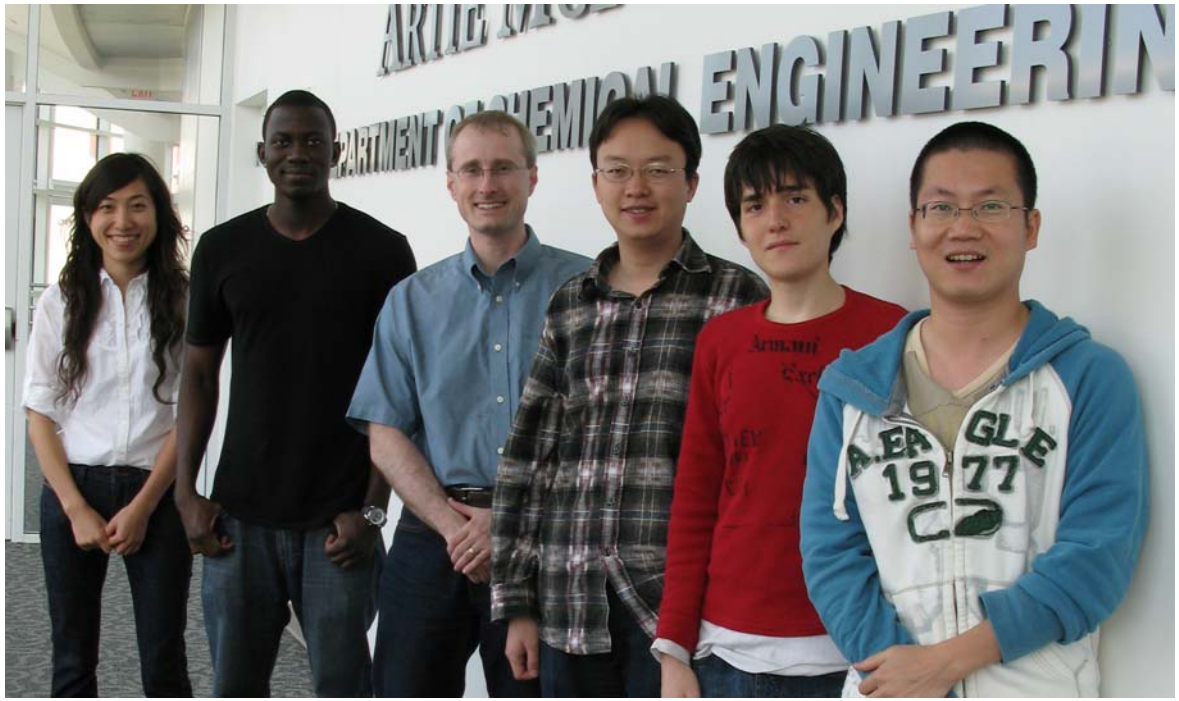
- A kinetic model for signal transduction of hepatocytes stimulated by IL-6 has been refined
- SOCS3, SHP-2, and PP2 are important proteins affecting the signal transduction as shown by simulations with knockout cells
- Sensitivity analysis was performed to quantitatively verify the conclusions drawn from the simulations
- Sensitivity analysis revealed a strong degree of cross-talk between the JAK/STAT and the MAPK pathway
- Based upon sensitivity analysis, a feedback loop where ERKPP phosphorylates Sos has been implemented

Conclusions

- Presented a method for quantitatively determining transcription factor concentration profiles from green fluorescent protein reporter systems
- Image analysis method has been investigated to obtain fluorescent intensity from images
- A model describing GFP dynamics is developed to describe the relationship between transcription factor and fluorescence intensity
- System inversion procedure has been developed to obtain quantitative data from fluorescence intensity profiles
- Updated model is in good agreement with data derived from GFP reporter systems

Acknowledgments

- Research Group
  - Obanifemi Aluko
  - Yunfei Chu
  - Bernardo Cunha
  - Jacky Huang
  - Cheryl Qu
- Collaborators
  - Arul Jayaraman
  - Fatih Senocak
- Funding Agencies
  - ACS-PRF, NSF, DOE



Thank you for your attention!

Questions ?